

Final Report

**An Interactive Environment for the Analysis of Large
Earth Observation and Model Data Sets**

Final
80
2000
33053
p 10

PI Name: Kenneth P. Bowman*Award Number:* NASA/USRA Subcont. 5555-04*PI Institution:* Texas A&M University*Report for Period:* June 5, 1991 — June 4, 1994*PI Address:* Climate System Research Prog.*Date of Report:* August 15, 1994

Department of Meteorology

Texas A&M University

College Station, TX 77845-3150

The purpose of this project was to develop interactive software for analyzing and visualizing large multi-dimensional geophysical data sets. Efforts under this contract resulted in the Envision software system. Envision binaries and source code can be obtained from either of the servers:

<ftp://vista.atmos.uiuc.edu/pub/envision/>

<ftp://csrp.tamu.edu/pub/envision/>.

Envision is in the public domain. Links to the Envision software can also be found in the Software Support Library online server. Detailed descriptions of Envision capabilities and operations can be found in the User's Guide and Reference Manuals distributed with the Envision software.

The Envision distribution includes a user manual, installation instructions, sample data sets, and connections to a variety of data display tools, including an IDL contouring module, an IDL line-graphing module, and NCSA Collage, and XDataSlice. The display tools available have been enhanced with mapping tools for contouring data on map projections (through the IDL interface) and with slicer options for the NCSA display tools.

Envision has a client-server architecture, where separate processes transmit and receive messages and data. Data transmission is accomplished using the NCSA DTM library. Envision was designed this way to make it easier for users to connect their own software packages to Envision. At present, there are two main components to Envision: one is the Data Manager, the other is the User Interface. The Data Manager acts as a data server. It reads and writes to your data files, keeps track of what data is where, and passes data to and from client processes. The User Interface is one client process. You can run the Data Manager without the User Interface, since other clients may connect with it, but it doesn't make sense to run the User Interface without a Data Manager. A graphics program

that plots data from your data files could be another client process. The Data Manager can talk to many clients simultaneously. This allows, for example, two users at different workstations to interact while browsing the same data sets.

Envision has online context-sensitive help, which uses the html version of the User's Guide. This help capability requires NCSA Mosaic, a free program used to access the World Wide Web.

Envision includes tools for managing large multi-file data sets. The data sets can be stored in one of two portable, self-describing data formats: netCDF and HDF. Envision can open any netCDF or HDF file and display the metadata in the User Interface.

The Envision User Interface displays metadata. Metadata is data about data. That is, it displays information about the data stored in your netCDF and HDF files, such as the variables in the file, their names and dimensions, any attributes stored in the file, such as units or scale factors, etc. The User Interface also enables the user to perform editing, adding, deleting, rearranging, and hiding of metadata. The User Interface can also generate a variety of metadata reports for a data set. Envision manages data in multiple file data sets by creating data objects such as logical dimensions and variables, and virtual dimensions. Using these, an entire data set can be viewed and used as if were all in a single file. The functionality of the Envision User Interface is complete, including attribute editing in project files. This allows users to create, change, delete, or restore dimension, variable, and global attributes within data files, including editing of dimension coordinate values. It is also possible to build arbitrary netCDF data files interactively from the User Interface.

Portability was one of the design goals of the Envision project. Through careful design and use of a number of industry-standard software tools, Envision will run on most common Unix workstations. Currently supported platforms are IBM RS6000's running AIX, HP 700 series computers running HP-UX, Sun workstations running SunOS, Silicon Graphics workstations running IRIX, and DEC Alpha workstations running OSF1. Source code is also provided for those who wish to port Envision to other platforms.

Systems supported in the Envision release:

IBM RS6000	AIX 3.2
H-P 700	HP-UX 9.0 (no IDL)
Sun Sparc	SunOS 4.1
Silicon Graphics	IRIX 5.2
DEC Alpha	OSF1 2.0

Publications

- Bowman, Kenneth P., John E. Walsh, and Robert B. Wilhelmson: An Interactive Environment for the Analysis of Large Earth Observation and Model Data Sets, extended abstract for the *American Meteorological Society Eighth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, to be held January 5–10, 1992, Atlanta, GA.
- Searight, Keith R., David P. Wojtowicz, Kenneth P. Bowman, Robert B. Wilhelmson, and John E. Walsh, 1993: Envision: A Collaborative Analysis and Display System for Large Geophysical Data Sets, *Ninth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, Anaheim, CA.
- Searight, Keith R., John E. Walsh, Kenneth P. Bowman, Sridhar Pathi, and Robert B. Wilhelmson, 1993: Envision: An Environment for Management and Visualization of Geophysical Data, , in *Developments and Applications of Data Handling and Visualization Tools to Space and Atmospheric Sciences*, eds. E.P. Szuszczeicz and J. Bredekamp, American Geophysical Union, Washington, DC (in press).
- Searight, Keith R., Xin Tao, William L. Chapman, John E. Walsh, Kenneth P. Bowman, and Robert B. Wilhelmson, 1994, A GCM Comparison Study Using Envision, *Proceedings of the Tenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, Nashville, TN, pp. 372–377.

This project developed the Envision client-server software system. Envision is software for interactive analysis and visualization of large, multi-dimensional, geophysical data sets on Unix workstations. Envision software (binaries and source code) and documentation can be obtained from either of the servers:

<ftp://vista.atmos.uiuc.edu/pub/envision/>

<ftp://csrp.tamu.edu/pub/envision/>.

Envision is in the public domain. Detailed descriptions of Envision capabilities and operations can be found in the User's Guide and Reference Manuals distributed with the Envision software.

Systems supported in the Envision release:

IBM RS6000

AIX 3.2

H-P 700

HP-UX 9.0 (no IDL tools)

Sun Sparc

SunOS 4.1

Silicon Graphics

IRIX 5.2

DEC Alpha

OSF1 2.0

Final Report

**An Interactive Environment for the Analysis of
Large
Earth Observation and Model Data Sets**

*Funded In Response to NRA-89-OSSA-21
Award NASA/USRA 5555-03*

June 5, 1991 - June 4, 1994

Principal Investigators:
Robert Wilhelmson
John Walsh

Department of Atmospheric Sciences
University of Illinois at Urbana-Champaign
105 S. Gregory Avenue
Urbana, IL 61801

Envision is an interactive environment that provides researchers in the earth sciences convenient ways to manage, browse, and visualize large observed or model data sets. Its main features are support for netCDF and HDF file formats, an easy to use X/Motif user interface, a client-server configuration, and portability to many UNIX workstations. The Envision package also provides new ways to view and change metadata in a set of data files. It permits a scientist to conveniently and efficiently manage large data sets consisting of many data files. It also provides links to popular visualization tools so that data can be quickly browsed. Envision is a public domain package, freely available to the scientific community.

Envision was developed to address the explosion in data volume that earth science researchers now work with, both in observational data sets and large model outputs. With large data sets, effective data management is critical. Another important consideration is the handling of metadata, or data about the data. Earth scientists must deal with data sets comprised of many files and need interactive, random access to any part of them. Data subsets should be easily and quickly displayed with a range of visualization options. Researchers must also be able to integrate their own visualization and analysis programs with a data management system. Envision's goal is to integrate the areas of data management, analysis, and visualization together in an interactive environment. The software's design is modular and distributed to allow for flexibility and extensibility.

Data sets used in Envision are stored as rectangular grids of values in n-dimensions in either netCDF or HDF formats. Envision is composed of a metadata browser and editor, a data management system, and a set of links to visualization and analysis tools. The key components run as separate processes in communication with each other. The data manager acts as a server, which connects to client processes, including the main user interface. External visualization and analysis tools act as clients and the data manager exchanges messages and data with them using the NCSA DTM library. Displays are made with NCSA's XDataSlice and Collage programs and RSI's IDL package, which are described below. Envision is extensible, since other tools could be added with a small amount of programming.

Data management is one of the key features of Envision. The data objects are "dimensions" (independent variables), "variables" (dependent variables), and "attributes" (descriptive information about variables and dimensions). The data manager organizes variables, dimensions, and attributes from a data set of related files into a "project." A large data set might be stored on a CD-ROM, with data files arranged in a hierarchy of directories representing years, months, and days. The data in these files is conceptually a single entity, but is split up for convenience of storage, or due to its collection over a period of time. In practice, data sets which are spread across a set of files can be unwieldy to manage. Problems arise if a user wishes to view a subset of the whole data set when the data values are in more than one file.

Management of variable and dimension objects that span multiple files is done in Envision by identifying and "merging" the objects that correspond in each of the files. For example, "Latitude" may be defined in each data file and is really the same dimension throughout the data set. A grouping of related dimensions or variables spanning multiple files is given a name and can then be used as if it were a single object in one file. Envision also creates dimensions implied by the division of data into files. A dimension, often time, may not be explicitly defined, but can be assumed when each file in a data set is a single step. The file name usually describes what time step is associated with that file's data. In Envision, this type of dimension can be named and used like an explicitly defined dimension. Envision's

management features for multiple file data sets let scientists access data without worrying about which files contain the values they need.

The user interface is the principal tool for working with a data set in an Envision project. The user interface is a point-and-click graphical environment and is an easy to use, intuitive way to browse both metadata and data. The metadata describes data values and includes information such as collection or generation of the data and its subsequent manipulation. The main part of the user interface is a table of dimensions and variables, as shown in Figure 1. At the top of the window is a menu bar, which has a series of pull-down menus for working with projects, data files, the appearance of the table, metadata access, and visualization. In the table display, dimensions are represented as columns and variables are shown as rows. Raised buttons indicate which dimensions are used by which variables. The units attributes, if defined, are displayed along with the variable and dimension names. The variables and dimensions in the table may all be from a single file or could span any number of files.

Project	Datafile	View	Variable	Dimension	Attribute	Display
	X	Y		Time months	Latitude degrees (N)	Longitude degrees (E)
Ozone relative units						
Temp degrees C						
Pressure millibars						

Figure 1. An example of Envision's main user interface.

One of the key uses of the Envision is to get access to the metadata associated with a project. In the past, metadata could be examined only by dumping the contents of a data file using a utility such as *ncdump*. Envision makes metadata browsing more convenient by interactively generating reports that describe variables, dimensions, attributes, and the data files they reside in. Metadata may also be deleted, changed, or added to using pull-down menu options. There are a variety of ways these applications can be used. One purpose is to make variable or dimension names more descriptive. For example, a variable named T could be changed to Temperature. If the units attribute were not defined for a dimension or variable, it could be added. If the metadata contained errors, these could be quickly corrected interactively.

As was mentioned previously, a data set may be on a read-only medium, such as a CD-ROM. In this situation, it is clearly not possible to modify anything in the data files. Since Envision stores metadata in a separate project file, metadata on read-only media can still be edited or added to. Even when users are able to change the data files, they may prefer to avoid the overhead of rewriting them. Envision also provides the option to write metadata

changes back to the data files and supports the creation of new data files with variables using subsets of dimension ranges from a project.

Envision does not do any visualization itself, but acts as a data server for visualization and analysis tools. Creating a display in Envision involves a few steps. First, the variable of interest is identified in the user interface. The dimensions to be used as the display axes are selected by clicking on the raised buttons in the variable's row. For example, to create a 2-D visualization, X and Y axes would be picked. A display program, such as Collage or IDL, is chosen from a pull-down menu. At this point, an "options" interface pops up. This interface allows the user to restrict the ranges of the dimensions and to set options for the visualization display. When the user requests a display, the data values are fetched from the data files, assembled in the proper order by the data manager, and sent directly to the visualization program. Other ranges or options can be set again and new displays created, without going back to the main user interface.

Two freely available visualization programs used with Envision are Collage and XDataSlice, from the National Center for Supercomputing Applications (NCSA). Both are raster image viewers that can display single frames or animations. Another program that Envision works with is Interactive Data Language (IDL), a product of Research Systems, Inc. IDL is an interactive environment used to perform calculations on and visualizations of scientific data. Modules written in IDL are used with Envision to create 1-D, 2-D contour, and 3-D surface plots. IDL is also used to generate geographic displays with continent outlines and various map projections. Figure 2 is an example of global climate model data plotted using the latter options.

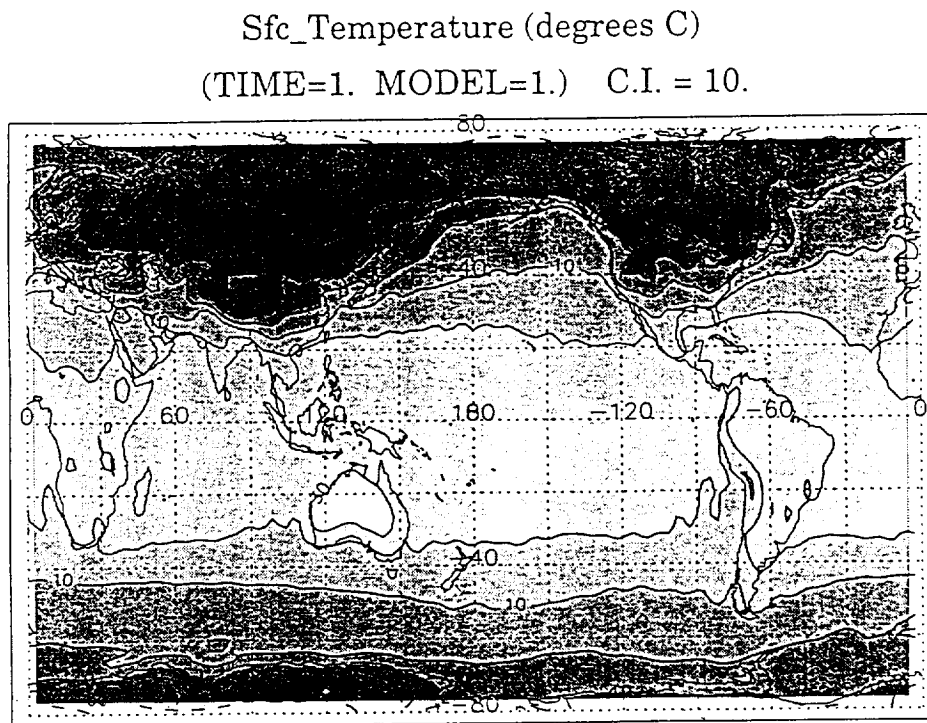


Figure 2. Geographical visualization of GCM data using IDL. The display shows surface temperature using filled contours dashed below zero degrees and a cylindrical projection. The data is from the ECMWF T42L19 AMIP 10 year simulation (1979-1988) for January 1979.

Envision runs on most popular UNIX workstations, including the IBM RS6000 series running AIX 3.2, the Sun Sparc running SunOS 4.1, the HP 700 series running HP-UX 9.0 (IDL link unavailable), Silicon Graphics workstations running IRIX 5.2, and the DEC Alpha running OSF1 2.0. No special hardware is needed. Source code is provided for those who wish to port Envision to other platforms. Envision is available by anonymous FTP at vista.atmos.uiuc.edu or csrp.tamu.edu in the subdirectory /pub/envision. Source, binaries, documentation, and sample data sets are provided, along with instructions for obtaining netCDF, HDF, Collage, XDataSlice, and IDL. Envision has online context-sensitive help, which uses the html version of the User's Guide. This help capability requires NCSA Mosaic, available from NCSA via the World Wide Web. The Envision distribution includes a user manual, installation instructions, sample data sets, and connections to a variety of data display tools, including an IDL contouring module, an IDL line-graphing module, NCSA Collage, and NCSA XDataSlice. The display tools available have been enhanced with mapping tools for contouring data on map projections (through the IDL interface) and with slicer options for both Collage and XDataSlice.

Users of the World Wide Web (WWW) and NCSA Mosaic can browse a hypermedia document containing text and graphics about Envision by opening URL <http://redrock.ncsa.uiuc.edu/Envision/envision.html>. There is also a WWW server at University of Colorado's Software Support Laboratory (SSL) which contains information about NASA's Applied Information Systems Research Program (AISRP) projects and related information. This can be accessed by opening URL http://sslabor.colorado.edu:2222/ssl/ssl_homepage.html in Mosaic. The hypermedia information there includes descriptive text, images, and links to anonymous FTP sites for obtaining software.

Publications

- Bowman, Kenneth P., John E. Walsh, and Robert B. Wilhelmson: An Interactive Environment for the Analysis of Large Earth Observation and Model Data Sets, extended abstract for the American Meteorological Society Eighth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, to be held January 5-10, 1992, Atlanta, GA.
- Searight, Keith R., David P. Wojtowicz, Kenneth P. Bowman, Robert B. Wilhelmson, and John E. Walsh, 1993: Envision: A Collaborative Analysis and Display System for Large Geophysical Data Sets, Ninth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, American Meteorological Society, Anaheim, CA.
- Searight, Keith R., John E. Walsh, Kenneth P. Bowman, Sridhar Pathi, and Robert B. Wilhelmson, 1993: Envision: An Environment for Management and Visualization of Geophysical Data, , in *Developments and Applications of Data Handling and Visualization Tools to Space and Atmospheric Sciences*, eds. E.P. Szuszczewicz and J. Bredekamp, American Geophysical Union, Washington, DC (in press).
- Searight, Keith R., Xin Tao, William L. Chapman, John E. Walsh, Kenneth P. Bowman, and Robert B. Wilhelmson, 1994, A GCM Comparison Study Using Envision, *Proceedings of the Tenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, Nashville, TN, pp. 372-377.
- Searight, Keith R., and J. E. Walsh, 1994: Envision: An Environment for Management and Visualization of Geophysical Data, *NASA Science Information Systems Newsletter*, NASA Office of Space Science, Issue 31, February, 1994, p. 3-6.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 1994	3. REPORT TYPE AND DATES COVERED Contractor Report	
4. TITLE AND SUBTITLE An Interactive Environment for the Analysis of Large Earth Observation and Model Data Sets			5. FUNDING NUMBERS 930	
6. AUTHOR(S) Kenneth P. Bowman, John E. Walsh, Robert B. Wilhelmson				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Texas A&M University & Department of Atmospheric Sciences College Station, TX 77843 University of Illinois 105 South Gregory Urbana, IL 61801-3070			8. PERFORMING ORGANIZATION REPORT NUMBER NAS5-32337 5555-03 5555-04	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration - OSSA Washington, D.C. 20546-0001 Universities Space Research Association 10227 Wincopin Circle, Suite 212 Columbia, MD 21044			10. SPONSORING/MONITORING AGENCY REPORT NUMBER CR-189381	
11. SUPPLEMENTARY NOTES When this project began, all three of the investigators were at the University of Illinois. Dr. Bowman subsequently moved to Texas A&M University. Ftp addresses are provided below for both. Technical Monitor: J. Hollis, Code 930				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified-Unlimited Subject Category 82 Report is available from the NASA Center for AeroSpace Information, 800 Elkridge Landing Road, Linthicum Heights, MD 21090; (301) 621-0390.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Envision is an interactive environment that provides researchers in the Earth sciences convenient ways to manage, browse, and visualize large observed or model data sets. Its main features are support for the netCDF and HDF file formats, an easy to use X/Motif user interface, a client-server configuration, and portability to many UNIX workstations. The Envision package also provides new ways to view and change metadata in a set of data files. It permits a scientist to conveniently and efficiently manage large data sets consisting of many data files. It also provides links to popular visualization tools so that data can be quickly browsed. Envision is a public domain package, freely available to the scientific community. Envision software (binaries and source code) and documentation can be obtained from either of these servers: ftp://vista.atmos.uiuc.edu/pub/envision/ and ftp://csr.p.tamu.edu/pub/envision/. Detailed descriptions of Envision capabilities and operations can be found in the User's Guide and Reference Manuals distributed with Envision software.				
14. SUBJECT TERMS user interface atmospheric sciences visualization data management GEMPAK data analysis Envision			15. NUMBER OF PAGES 8	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

ORIGINAL PAGE IS
OF POOR QUALITY